



Gegevens de Baas

Eric Brouwer

Gegevensmanagement, 4 april 2019 te Amersfoort

1

“Zijn gegevens de baas?”

of

“Zijn wij de gegevens de baas?”

Een hoge kwaliteit van de gegevenshuishouding is nodig voor goede dienstverlening aan burgers en bedrijven én voor efficiënt samenwerkende overheidsorganisaties.

De processen van het verzamelen, registreren en gebruiken van de gegevens zijn van grote invloed daarop.

Wie bepaalt dan eigenlijk wanneer gegevensmanagement goed is?

En welke norm, waardenstelsel wordt daarbij gehanteerd?

AP17 gegevens beschrijven en modelleren

thema Gegevens management

thema Semantiek

www.NORAonline.nl

3

De Nederlandse Overheid Referentie Architectuur (NORA) is een door het Kabinet (in 2009) vastgesteld afsprakenstelsel voor het ontwerp van de dienstverlening van de overheid: te gebruiken door alle overheidsorganisaties !

De NORA omvat richtinggevend principes en handvatten voor architecten, zodat zij een goed ontwerp kunnen maken van de overheidsdiensten.

AP17 is het principe dat aangeeft dat als organisaties met elkaar samenwerken, dat dan afspraken worden gemaakt over de daarbij te gebruiken gegevens.

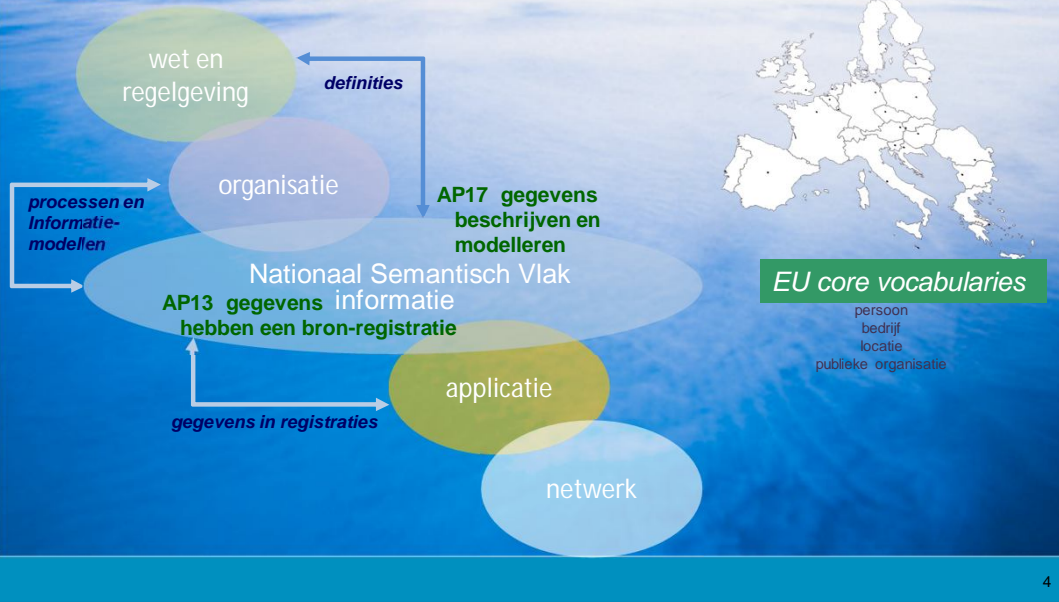
Maar toen we in 2012 wilden weten welke afspraken dan waren gemaakt, was er nergens een overzicht te vinden !

We hebben toen via Google diverse gegevenswoordenboeken opgespoord, de beheerders ervan gevraagd met ons samen het thema Semantiek op te zetten en ook AP17 uit te werken, zodat we meer eenduidige beschrijvingen gaan krijgen. En in 2018 is AP17 opnieuw geactualiseerd op basis van goede ervaringen met de modelleringswijze van de Omgevingswet.

Het thema Gegevensmanagement is vanaf 2018 in de NORA opgenomen: "overerfd" vanuit de Gemma.

Het NSV, samen met het MIM, wordt daarbij gehanteerd als basis om een gemeenschappelijke taal te ontwikkelen om gegevenswoordenboeken (en vergelijkbare termen zoals informatiemodellen) te beschrijven.

Gegevens zijn de bron van onze informatievoorziening



De NORA omvat een vijflaagsmodel (5 verschillende views) waarmee architecten een vraagstuk uiteen kunnen rafelen in voor niet-architecten begrijpelijke taal en oplossingsrichtingen.

De stakeholders kunnen daardoor besluiten nemen op basis van de door de architect aangegeven impact op maatschappelijk, organisatorisch en technisch niveau.

Een van die 5 views gaat over informatie, waaronder semantiek, gegevensmanagement en data op het web.

Al onze kennis en ervaring wordt vrijelijk gedeeld (onder Common Creative copyrights) en is daartoe beschikbaar op het internet !

Het gedachtegoed over een Nationaal Semantisch Vlak is begin 2013 ontstaan. In 2016 is het inhoudelijk verder uitgewerkt en daarnaast breidt de context zich langzaam uit richting Europa. Zie http://www.noraonline.nl/wiki/Nationaal_Semantisch_Vlak

Onze aanpak is gericht op de semantiek (de begrippen en de definities) om zo voorlopig weg te blijven uit de never-ending discussions over privacy of autorisaties.

De bron van veel definities is te vinden in de Nederlandse Wet- en Regelgeving: www.wetten.overheid.nl

Doorgaans wordt voor elk systeem, elke organisatie en elk domein, een afzonderlijk gegevenswoordenboek en/of gegevensmodel gehanteerd.

Het Nationaal Semantisch Vlak geeft een overzicht van bij ons bekende gegevensbeschrijvingen en plaatsen waar meer informatie te verkrijgen is.

Het idee van een NSV omzetten naar acties die meerwaarde in de dienstverlening geven, vraagt om een aanpak die voorkomt dat het als een academische exercitie wordt gevoeld.

Wat zijn dan de juiste momenten waarop actie nodig is?

Zie daarvoor de presentatie

https://www.noraonline.nl/images/noraonline/e/e3/WIV12062015_NationaalSemantischVlak.pdf

Daarbij is het van belang de (bestuurlijke) ontwikkelingen van het "Gegevenslandschap van onze Samenleving" te volgen, zie <https://www.digitaleoverheid.nl/dossiers/gegevenslandschap/>

Er is veel tijd te winnen door her-gebruik van gegevensbeschrijvingen en gegevensmodellen

5

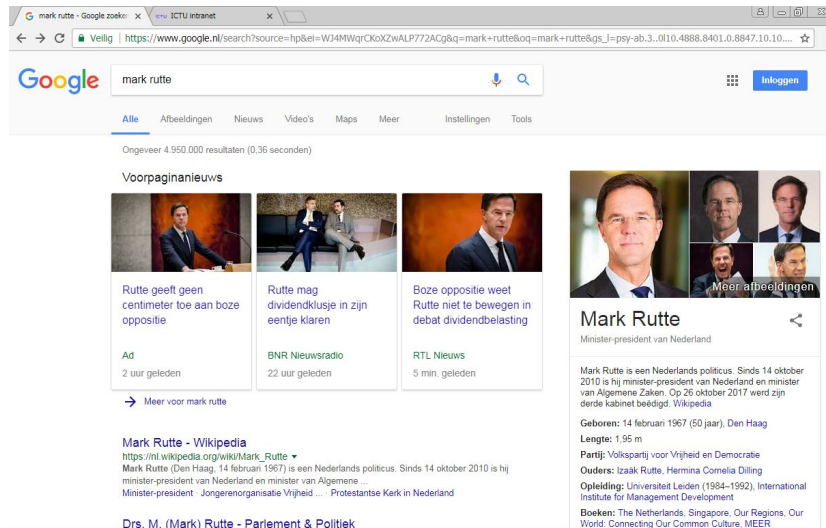
Iedere organisatie besteedt veel tijd aan het definiëren en modelleren van de gegevens waarmee ze te maken heeft.

Stel je voor dat je die definities eenvoudig van elkaar zou kunnen her-gebruiken.

Niet via copy-paste, maar door internet-linkjes en real-time presentatie.

**Via semantische wiki's zoals de NORA
is dat mogelijk !**

Dit doet Linked Data ...



Als je Google't, dan wordt het resultaat bijeengebracht vanuit allerlei servers over de gehele wereld.

Bekende meta-data worden automatisch gelinkt.

Daarmee ontstaat vanzelf een afdoende bruikbare definitie.

En Artificial Intelligence in combinatie met zoekmachines zal er voor gaan zorgen dat die meta-data en "definities" meer en meer contextueel verbijzonderd worden.

Bijvoorbeeld afhankelijk vanaf welke locatie je zoekt of wat je eerder hebt gezocht.

Dat is wat we willen bereiken.

Met dit verschil, dat we het via Linked Open Data beschikbaar willen hebben, op een open-source eco-systeem c.q. datanutsvoorziening.

En dus niet dat Google of een andere private partij allerlei aan ons gerelateerde gegevens zomaar in zijn eigen databases cachet (dus meta-data en brondata) en daar van alles mee gaat doen ...

Er is veel tijd te winnen door her-gebruik van gegevens via Linked Data

8

Linked Data lijkt een ontwikkeling te zijn die ons in staat stelt om grote stappen te maken qua - wereldwijde ! – afstemming van terminologie, semantiek en gegevens. Niet meer kopiëren, maar verwijzen !

De vraag is daarbij zelfs of we nog wel definities nodig hebben ...

Definities zijn primair voor mensen, want in de praktijk hebben die vooral behoefte aan een vertaling naar “klare taal”.

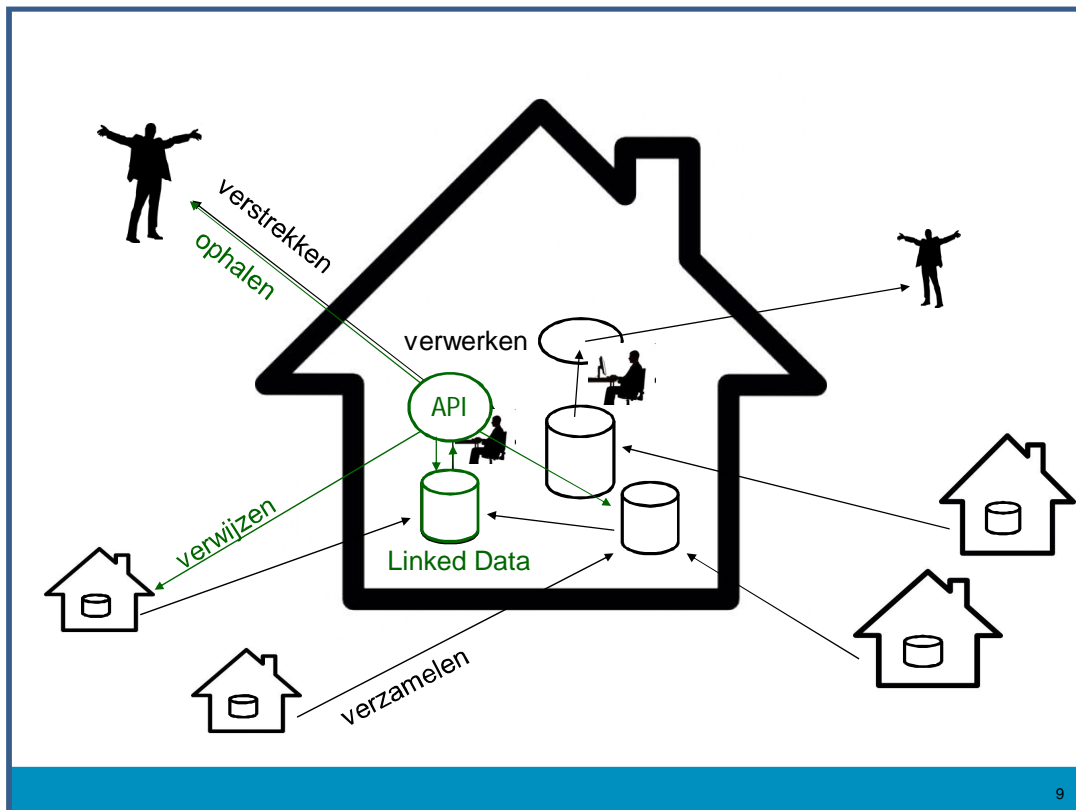
Computers kunnen op basis van structuren al een heleboel semantiek herkennen. Een voorbeeld uit de BRK: een Zakelijk recht *is een* Recht *dat betrekking heeft op* een Zaak.

Computers kunnen gegevens direct “aanwijzen” (linken), zonder gebruik te maken van een definitie die we als mensen hanteren. Ze verwijzen direct naar waarden in databases, dus naar de betreffende informatieobjecten. Dat komt overeen met een concrete tafel, auto of een persoon uit de werkelijkheid. En zelfs inclusief allerlei meta-gegevens, die via het internet en met business-rules geautomatiseerd worden gelinkt.

Hoe zit dat dan?

Wie toe is aan een nieuw perspectief op gegevens:

<https://www.overheidincontact.nl/een-nieuwe-visie-ook-met-gegevens-werken-vanuit-de-bedoeling/>



Traditioneel verzamelt elke organisatie de relevante gegevens en slaat die op in eigen registraties om van daar uit diensten te verlenen.

Relaties tussen gegevens in de databases worden gelegd via zogenaamde sleutels.

Deze werkwijze is logisch: hierdoor zijn de taken zelfstandig uit te voeren en is de continuïteit van dienstverlening redelijk goed gegarandeerd.

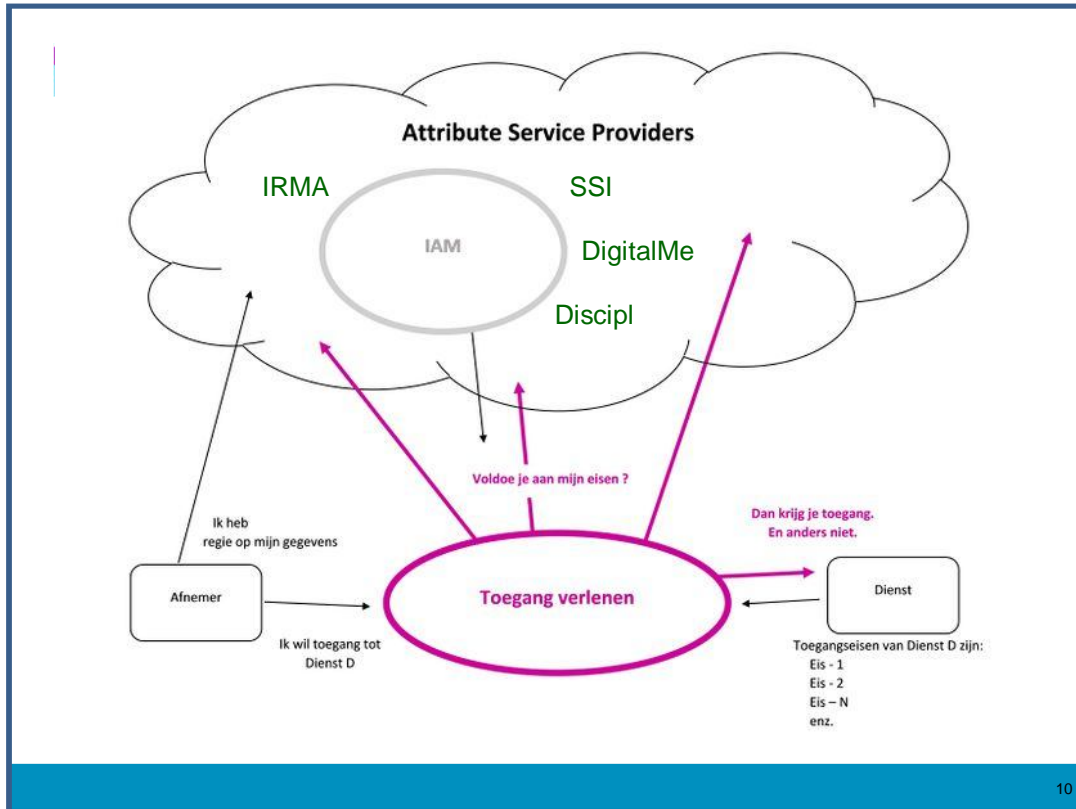
Echter, het betekent ook veel dubbel werk, gegevens zijn dan lang niet altijd actueel en vooral kan het leiden tot lange doorlooptijden als je in ketens samenwerkt.

Door de technologische ontwikkeling van Linked Data hoeft dat verzamelen in het "eigen huis" niet meer: via het gebruik van Application Programming Interfaces (API's) zijn gegevens real-time direct vanuit registraties bij andere organisaties te gebruiken. Wereldwijd !

De sleutels (relaties tussen de gegevens) zijn daarbij uri's geworden die naar de bron-gegevens verwijzen.

Dit vereist natuurlijk wel dat registraties zijn opgezet met Linked Data, hetgeen doorgaans nog niet het geval is.

Daar kan wel snel een business case voor zijn: gegevens kunnen aan derden worden verstrekt zonder tussenkomst van menselijk handelen, waardoor verstrekingskosten substantieel kunnen verminderen.



Hier zien we wat in de nabije toekomst gaat gebeuren.

Als je een dienst digitaal wil afnemen, dan komt een dialoog tot stand om te kijken of je aan de gestelde eisen van de dienst aanbieder voldoet.

We gaan dan niet meer onze naam, geboortedatum e.d. op een (e)formulier invullen, maar we laten de computers in de cloud het werk voor ons doen: ze gaan straks “verifieerbare claims” met elkaar uitwisselen, via wereldwijde op open source gebaseerde datanutsvoorzieningen (zoals het internet dat we nu kennen, maar dan uitgebreid met extra functionaliteit voor veiligheid en privacy).

Zo’n claim is dan bijvoorbeeld dat je ouder bent dan 18 jaar (ja/nee).

Hoe je heet hoeft de dienst aanbieder misschien niet eens meer te weten: de persoon die aan de eisen voldoet krijgt toegang tot de dienst.

Je moet dan natuurlijk wel zelf aangeven waar die computers die claims kunnen verkrijgen. En mogelijk moet je daarbij ook aantonen wie je bent (bv door in te loggen met DigiD en via de BRP te laten verklaren dat je ouder bent dan 18 jaar).

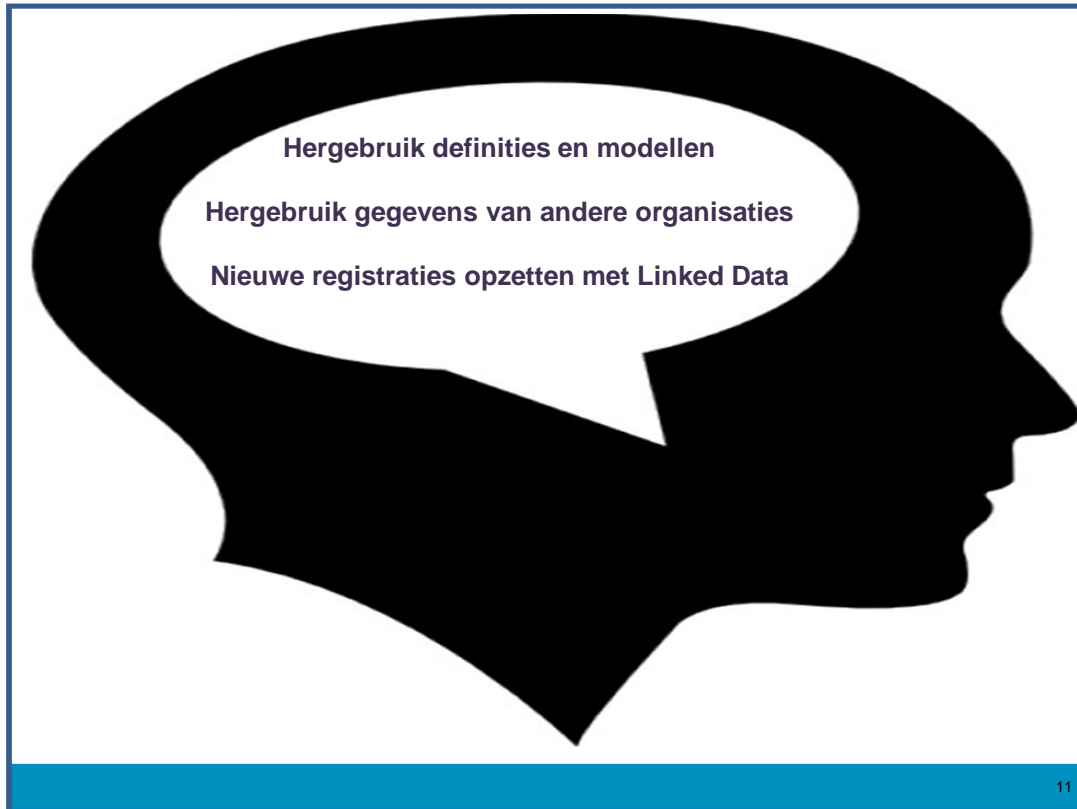
Ofwel: je zult regie (moeten) voeren 😊

Onderdelen daarvan worden in Nederland al zichtbaar via:

- IRMA <https://privacybydesign.foundation/irma/>
- DigitalMe <https://digital-me.nl/>
- SSI <https://sovrin.org/>
- Discipl <https://discipl.org/>

Bron: [https://www.noraonline.nl/wiki/Identity %26 Access Management](https://www.noraonline.nl/wiki/Identity_%26_Access_Management)

En zo krijgen we wellicht toch regie op “onze” gegevens !



Deze 3 boodschappen willen we graag meegeven aan een ieder die Gegevensmanagement een warm hart toedraagt.



Heb je vragen of aanvullingen?

Stuur dan even een mailtje !